UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL

Selection of single domain antibodies against the CD20 receptor. A new treatment with potential anti-tumor properties for B-cell malignancies

Mestrado em Biologia Molecular e Genética

Ana Filipa Santos André

Dissertação orientada por:

Doutor Frederico Aires da Silva (FMV-UL) e Prof. Doutor Rui Malhó (FCUL)

Versão Pública

2016

Acknowledgements

"Recomeça… Se puderes, sem angustia e sem pressa. E os passos que deres, nesse caminho duro do futuro, dá-os em liberdade. Enquanto não alcances não descanses. De nenhum fruto queiras só metade."

Miguel Torga

Em todas as etapas das nossas vidas, cruzamo-nos com pessoas excecionais e outras nem tanto, mas é isso que nos torna especiais e faz com que toda esta caminhada não seja solitária. Por isso, chegou o momento de agradecer a todas essas pessoas excecionais com quem me cruzei durante esta etapa.

Quero começar por agradecer ao meu orientador, o Doutor Frederico Aires da Silva, por me ter apresentado este projeto e por me transmitir os seus conhecimentos e toda a sua paixão por este grande mundo que é o dos "Anticorpos". Muito Obrigada pelo voto de confiança, por acreditar nas minhas capacidades e me incentivar a ir sempre mais longe.

Ao meu orientador na FCUL, o Prof. Rui Malhó, pelo acompanhamento nesta etapa, pela disponibilidade demonstrada e pela partilha de conhecimentos.

Às minhas colegas de laboratório Joana Dias e Sandra Aguiar, por me terem recebido da melhor maneira, por toda a vossa paciência para partilhar conhecimentos e pelos vossos conselhos que muito me ensinaram. Tornaram os meus dias mais alegres e ajudaram a transformar os obstáculos em pequenas vitórias.

A todas as pessoas com que me cruzei na FMV, na FCUL e na Technophage, e que de alguma forma contribuíram não só para a realização deste trabalho mas também para o meu bem-estar, mesmo que apenas com um "Bom Dia" todas as manhãs.

"Happiness is real friends, who care."

À minha amiga Ana Cachucho que me tem acompanhado nesta aventura da ciência desde o inicio. Obrigada por toda a motivação, por não me deixares desistir e por acreditares em mim, mesmo quando eu mesma duvidei. Há pessoas que sabemos que irão permanecer e sei que serás uma delas.

À minha amiga e eterna colega de casa Carina Maranga, (passados dois anos, continuas a ser a minha colega de casa preferida). Mesmo a uns bons kms de distância estaremos sempre perto. Obrigada por todo o teu apoio, incentivo, mas sobretudo pela tua amizade, que sei que é verdadeira.

A todos os meus amigos e pessoas que são essenciais na minha vida, que mesmo distantes me fazem sentir acompanhada, sem nunca questionar as minhas escolhas.

"The greatest happiness is family happiness".

Ao meu Filipe, por tudo. Pela ENORME ajuda nesta etapa, por toda a paciência para os momentos menos bons e acima de tudo por continuares sempre aqui. Obrigada talvez seja pouco, mas por vezes as palavras são demasiado pobres para expressar tão grandes sentimentos. Sem ti nada seria igual. "Happiness is having you in my life".

À minha mãe, o meu exemplo de força e determinação. Obrigada pelo teu apoio incondicional, por me deixares sempre fazer as minhas escolhas e me indicares o melhor percurso a seguir. Será sempre um orgulho poder dizer que sou tua filha. "Happiness is a mom like you".

À minha irmã, a minha companheira de todas as aventuras, aquela pessoa que sei que estará sempre ao meu lado, independentemente de tudo o resto. Obrigada por todas as palavras, por toda a força e por seres assim, exatamente como és. Sei que um dia serás tu também e, como sempre, estarei aqui com o imenso orgulho que tenho em ti. "Happiness is unconditional sisterly love".

Aos meus avós e ao meu tio, por toda a preocupação, por todo o apoio e por tudo o que fazem por mim. Obrigada por estarem sempre presentes!

Aos meus tios e primo, que durante estes dois anos nunca me deixaram desamparada nesta nova cidade que tive de descobrir.

Abstract

Lymphoma is the third most common neoplasia in the world. Within lymphomas, non-Hodgkin lymphoma (NHL) is the most common affecting mainly B cells.

For more than three decades, chemotherapy and radiotherapy have been the only treatments available, but since the discovery of Rituximab, a new era of lymphoma therapy was inaugurated, once it was the first monoclonal antibody approved against the CD20 receptor.

The CD20 receptor is an antigen expressed on the B-cells surface, and it is present nearly in all of its maturational development, being absent only in the stages of the pro-B lymphocyte and plasma cells. Furthermore, the CD20 receptor is also present on >90% of the B-cell NHL. These characteristics make this receptor an ideal target for immunotherapy.

Besides the success of Rituximab, various mechanisms of resistance have been developed by the tumoral cells against this antibody, such as the decrease of the CD20 expression on B-cells, immunogenicity, structural changes affecting the binding region of the CD20 antibody or alterations in the cell membrane. New anti-CD20 antibodies have been developed to overcome the disadvantages of Rituximab, such as Ofatumumab and Obinutuzumab, presenting a different immunogenicity and binding to the different epitopes on the target with a different affinity. Apart from the progress in the lymphoma therapy, a significant population of patients still succumbs to this disease, as tumoral cells are always changing, making the search for better antibodies a never-ending process. Thus, the aim of this project consisted in the development and characterization of a new antibody against the CD20 receptor. To achieve this goal, the potential of rabbit derived single-domain antibodies (sdAbs) as therapeutic molecules were explored. For that, one rabbit was immunized with the CD20 receptor and an immune VH and VL sdAb library was generated. Then, a subtractive cell phage display screening was used for antibody selection. This approach allowed a specific selection of one sdAb in the VL format against the CD20 receptor in cells. In summary, the strategy explored in the present project resulted in an antibody that, according to its characteristics, could be a promising candidate in the treatment of NHL and other B-cell malignancies.

Keywords:

CD20; Lymphoma; Immunotherapy; Single domain antibodies; Phage Display

Resumo alargado

Os linfomas são a terceira neoplasia mais comum no mundo, sendo os linfomas não-Hodgkin os mais frequentes. Dentro dos linfomas não-Hodgkin, cerca de 85-90% são de células B. Os linfomas desenvolvem-se maioritariamente a partir dos nódulos linfáticos, no entanto podem formar-se a partir de qualquer outro tecido.

Durante mais de três décadas, os únicos tratamentos disponíveis para esta doença eram a quimioterapia e a radioterapia. No entanto, com a descoberta do Rituximab, um anticorpo específico para o recetor CD20 presente nos linfomas de células B, começou uma nova era no campo das imunoterapias.

O anticorpo Rituximab, uma vez que foi o primeiro a ser descoberto, faz parte da primeira geração de anticorpos anti-CD20. Este é um anticorpo quimérico, uma vez que é constituído por um IgG1 glicosilado composto por uma região constante kappa humana e regiões variáveis leves e pesadas murinas. Este anticorpo atua essencialmente através de três principais mecanismos de ação: citotoxicidade mediada por anticorpos (ADCC); citotoxicidade dependente do complemento (CDC) e indução da apoptose.

O recetor CD20 é uma molécula expressa na superfície das células B e está presente na maioria das fases de maturação destas células, com exceção da fase de linfócitos pró-B e nas células plasmáticas. Para além disso, está presente em mais de 90% dos linfomas não-Hodgkin de células B. No entanto, apesar da sua função ainda não ser totalmente conhecida, pensa-se que possa estar envolvido no fluxo de cálcio. Este recetor é uma fosfoproteína não glicosilada com 4 domínios membranares: os domínios N e C terminal que são intracitoplasmáticos e 2 loops adicionais que são extracelulares. Devido a estas suas características, o recetor CD20 é considerado um bom alvo para imunoterapias, nomeadamente terapias para linfomas não-Hodgkin de células B.

Apesar do sucesso do Rituximab, têm vindo a ser descritos vários mecanismos de resistência das células tumorais contra este anticorpo, nomeadamente a diminuição da expressão de CD20, mudanças estruturais que afetam o local de ligação do anticorpo à molécula ou alterações na membrana da célula, o que diminuiu a sua eficácia. Este aspeto fez com que fosse necessária a contínua procura por novos anticorpos para colmatar as lacunas do Rituximab. Assim, ao longo dos anos, foram descobertos novos anticorpos específicos para CD20, como o Ofatumumab e o Obinutuzumab, que têm uma diferente imunogenicidade e reconhecem diferentes epítopos na molécula CD20 aos quais se ligam com diferentes afinidades, atuando assim com diferentes mecanismos de ação. Ainda assim, este tipo de doença continua a afetar imensas pessoas, muitas vezes não existindo um tratamento eficaz devido às múltiplas resistências das células tumorais. Tendo em conta este contexto e uma vez que os anticorpos disponíveis no mercado não

conseguem resolver estes problemas, a procura por novos anticorpos mais vantajosos continua. Atendendo a esta necessidade, o objetivo deste projeto é o desenvolvimento e caracterização de um anticorpo de pequeno domínio específico para CD20. Os anticorpos de pequenos domínios têm várias vantagens em relação aos IgGs completos, principalmente devido ao seu reduzido tamanho, que faz com que estes anticorpos tenham melhor acesso ao alvo na superfície da célula, para além disso podem ainda ser facilmente expressos em bactérias como proteína. Têm ainda a vantagem, como moléculas terapêuticas, de serem mais estáveis em circulação que os anticorpos completos. Estas características permitem uma administração de maiores quantidades por grama de produto, o que leva a um aumento significativo da potência por dose e na redução do custo de produção.

Para o desenvolvimento do projeto, inicialmente foi imunizado um coelho com o péptido CD20 e com células HEK 293T transfetadas com o recetor CD20. Antes de cada imunização, o soro era recolhido para ser avaliado através de ELISA. Ao dia 89, quando foi atingido um elevado título de anticorpos, procedeu-se à recolha do soro final e à eutanásia do coelho. Os ensaios de ELISA permitiram confirmar que o soro extraído após as imunizações estava a reconhecer não só as células Raji, que contêm à sua superfície CD20, mas também o péptido CD20, por outro lado o soro recolhido antes das imunizações não reconhecia nem as células Raji, nem o péptido. Após a eutanásia, foram recolhidos o baço e a medula que são os órgãos onde se encontra um maior número de células plasmáticas que produzem anticorpos. A partir destes órgãos, procedeu-se à extração de ARN usando o reagente Tri. A partir do ARN extraído foi possível sintetizar a primeira cadeia de ADN. O cADN sintetizado foi utilizado para a construção de bibliotecas imunes de pequenos domínios de anticorpos específicos para o recetor CD20, através da amplificação por PCR das famílias de cadeias variáveis leves e pesadas. Esses produtos resultantes da amplificação foram clonados no vetor pComb3x, que é necessário na técnica de phage display, resultando numa biblioteca imune com uma grande diversidade, ideal para este tipo de seleção. Em seguida, as bibliotecas resultantes foram selecionadas através da tecnologia de phage display que permite a seleção de anticorpos específicos para CD20 em células que contenham esse recetor. Esta tecnologia baseia-se na engenharia genética de bacteriófagos e em várias rondas de seleção contra o antigénio e propagação dos fagos. Resumidamente, os fagos vão conter à sua superfície o fenótipo correspondente ao genótipo encapsulado no seu interior que contem a sequência correspondente aos fragmentos de ADN obtidos através da construção das bibliotecas de anticorpos. Seguem-se várias rondas de seleção onde vão sendo eliminados os anticorpos-fagos que não se ligam ou que possuem uma fraca ligação às células que contêm o antigénio, neste caso o recetor CD20, através de várias lavagens. Estas rondas vão sendo repetidas até obtermos uma população de anticorpos específica para o alvo. Para este phage display, foram utilizados 3 tipos de células: células

Raji, que está descrito que expressam o CD20, pois são células de linfoma B; células HEK 293T que não expressam CD20 e por isso foram utilizadas para eliminar os anticorpos não específicos para este recetor; células Jurkat, que tal como as anteriores não possuem CD20, pois tratam-se de células T de leucemia aguda e que por isso foram também utilizadas para eliminar anticorpos não específicos. Esta seleção resultou num conjunto de anticorpos-fagos específicos para o recetor CD20, resultado que foi confirmado através de Western Blot. Após a seleção por phage display, foi necessário fazer uma seleção em larga escala para avaliar os anticorpos que melhor eram expressos e que melhor se ligavam ao alvo. Para esta avaliação da expressão foi necessário clonar o ADN resultante dos fagos selecionados por phage display, num outro vetor que permitisse a expressão de proteína, neste caso o vetor pT7-PL. Esta avaliação da ligação e expressão foi feita através de ensaios de ELISA com extrato proteico de células Raji e células Raji. Para além disso, após várias seleções os melhores 10 clones foram sequenciados para avaliar a homologia entre si e verificar os CDRs. Através do alinhamento da sequência de aminoácidos obtida a partir da sequenciação, verificou-se que 9 dos clones eram iguais, restando assim apenas 2 clones. A especificidade dos dois clones selecionados para o recetor CD20 foi avaliada através de Western Blot, sendo que um deles (anticorpo no formato VL) era específico para o recetor pretendido.

Estas metodologias levaram à seleção de um anticorpo de pequeno domínio específico para CD20, que devido às suas características únicas e diferentes dos que já existem no mercado poderá ser um promissor candidato para ser usado como agente terapêutico para linfomas de células B que expressem o recetor CD20.

Palavras- Chave:

CD20; Linfoma; Imunoterapia; Anticorpos de pequeno domínio; Phage Display

Index

List of figures and tables

List of Abbreviations

aa	Amino acids
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
ADCC	Antibody-Dependent Cell-mediated Cytotoxicity
Amp	Ampicillin
Anti-HA-HRP	Anti-Hemagglutinin-Horseradish Peroxidase
BM	Bone Marrow
BSA	Bovine Serum Albumin
C	Carboxylic terminal
CDC	Complement-Dependent Cytotoxicity
cDNA	complementary DNA
CDR	Complementary Determining Region
CH	Constant domain of heavy-chain
CL	Constant domain of light-chain
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic Acid
E. coli	*Escherichia coli*
ELISA	Enzyme- Linked Immunosorbent Assay
Fab	Fragment antigen binding
Fc	Constant Fragment
FDA	Food and Drug Administration
LB	Luria-Bertani Broth
mAbs	Monoclonal antibodies
N	Amino Terminal
NHL	non-Hodgkin Lymphoma
NK cells	Natural Killer
OD	Optical density
PBS	Phosphate Buffered Saline

PCR	Polymerase Chain Reaction
PEG-8000	Polyethylene glycol
pT7- PL	pT7-peptide leader
RNA	Ribonucleic Acid
RPM	Rotations per minute
RPMI	Roswell Park Memorial Institute Medium
RT	Room temperature
SB	Super Broth medium
scFv	Single-chain variable fragment
sdAb	Single Domain antibody
SOC	Super Optimal Broth
SP	Spleen
UV	Ultra violet
VH	Variable domain of heavy-chain
VL	Variable domain of light chain

1. Introduction

1.1. Cancer and Lymphoma

Cancer is the second most common cause of death in the United States, making this disease the major public health problem in the world, killing one in every four persons, remaining one of the most challenging diseases to treat.

Lymphoma, which is one kind of hematological cancer, is the third most common neoplasia in the world and is predominantly a cancer of B-lymphocytes.[1,2] Within lymphomas, non-Hodgkin lymphoma (NHL) is the most common hematologic malignancy and accounts for approximately 4% of all adult cancers in the United States.[1] NHL comprises a heterogeneous group of cancers: 85-90% from B lymphocytes and the remainder derives from T lymphocytes or NK cells. These malignancies usually develop in the lymph nodes, but can occur in almost any tissue.[3]

1.1.1. Current and emerging therapies for Cancer and Lymphoma

For more than three decades, the only therapies available for lymphoma and other kinds of cancer were chemotherapy and radiotherapy.[4] However, with the advances in the field of immunology, recombinant DNA technology and protein engineering, new therapies for cancer emerged over the past years. These therapies, known as immunotherapies, use part of the immune system to fight diseases, stimulating it to specifically attack cancer cells, or even using immune system proteins.

There are various types of immunotherapies, but monoclonal antibodies (mAbs) are one of the most important and successful strategies for treating patients.[5] Currently, about fifty monoclonal antibodies have been approved in the United States and Europe for treatment of a variety of diseases including cancer.[6,7]

Antibody therapy can function through mediating alterations in antigen or receptor function, modulating the immune system or delivering a specifying drug that is conjugated to an antibody that targets a specific antigen. Mechanisms of cell killing by antibodies can be summarized in direct action of the antibody; immune-mediated cell killing mechanisms and specific effects of an antibody on tumor vasculature and stroma. These mechanisms are outlined in Fig. 1.[5]

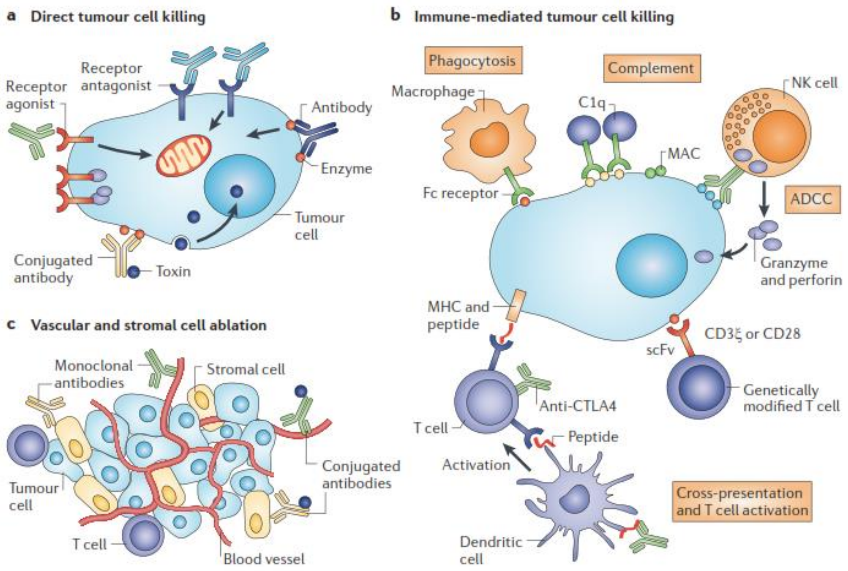


Figure 1 - Mechanisms of cell killing by antibodies. a) Direct tumor cell killing can be mediated by receptor agonist activity or antagonist activity leading both to apoptosis. It can also be mediated by a conjugated antibody used to deliver a payload to a tumor cell. b) immune-mediated tumor cell killing include mechanisms such as induction of phagocytosis, complement activation (CDC); antibody-dependent cellular cytotoxicity (ADCC); genetic modified T cells targeted to the tumor by single-chain variable fragments (scFv); T cells being activated by antibody-mediated cross-presentation of antigen to dendritic cells; and inhibition of T cell inhibitory receptor. c) vascular and stromal ablation can be induced by stromal cell inhibition; delivery of a toxin to stromal cells; and delivery of a toxin to the vasculature. Adapted from Scott *et al*, 2012.[5]

1.1.2. CD20 as therapeutic target

The MS4A1 gene located on chromosome 11q12.2 encodes the human CD20. This receptor is a non-glycosylated phosphoprotein with four membranes spanning domains, the N - and C – terminal domains, which are intracytoplasmic and two additional extracellular loops (Supplementary figure 1). Its functions are not fully understood, however, it is thought that they may be involved in the calcium flux.[8]

The CD20 receptor is expressed on the B-cells surface and is present during nearly all its maturational development, only being absent in the pro-B lymphocyte and plasma cells. Furthermore, CD20 is also

present on >90% of the B-cell NHL.[8,9] Therefore, these characteristics make this receptor an ideal target for immunotherapy.

The central challenge of immunotherapies has been the identification of suitable targets. Due to its characteristics, the CD20 receptor is one of the most promising targets for this approach.

For lymphoma treatment, Rituximab was the first monoclonal antibody against the CD20 receptor approved by the FDA, in 1997.[10] Since its approval, chemotherapy in combination with immunotherapy remains the standard treatment.[11]

Throughout the years, the discovery of new antibodies for the CD20 receptor contributed to the development of new therapeutics and the quality of life of patients. Nevertheless, tumoral cells are always changing, making the search for better antibodies a never-ending process.

1.2. Antibodies

1.2.1. Characteristics and proprieties of antibodies

Antibodies, also named immunoglobulins, are a family of glycoproteins produced in response to exposure to foreign structures, known as antigens. Nonetheless, antibodies display remarkable variability in the regions that bind antigens, possess a common structure of four polypeptide chains, consisting of two identical light chains and two identical heavy chains, which confer to each antibody at least two antigen-binding sites. Heavy and light chains are made of amino-terminal variable regions and carboxyl-terminal constant regions. In light chains, the variable and constant regions are composed of one VL and CL domain, while heavy chains are made of one variable domain (VH) and three constant domains (CH1, CH2 and CH3). The variable domains of light and heavy chains contain three hypervariable regions that are also known as complementary-determining regions (CDRs), as they can interact with the bound antigen. These regions differ in length and sequence among antibodies, which is why they are responsible for the specificity and affinity of the antibodies to the antigen. Light chains are bound to the heavy chains by a disulfide bond formed between cysteine residues and multiple non-covalent interactions. Heavy chains are also covalently bound by disulfide bonds. Between the CH1 and CH2 domains is located the hinge region, which is a flexible segment of the heavy chain that allows the two antigen-binding sites to operate independently. The antibody molecule forms a Y shape, meaning that the two identical heavy and light chains contribute to each arm of Y, while its base is composed by C-terminal domains of the heavy chain.

The two-paired heavy and light chains are named Fab (antigen-binding fragments), responsible for the antigen binding, while the C-terminal domains are known as constant fragment (Fc), responsible for the recruitment of the effector functions of the immune system.

In humans, it is possible to recognize five classes of immunoglobulins, namely IgG, IgA, IgM, IgD and IgE, which can differ in size, charge, amino acid sequence and carbohydrate content. IgG is the most predominant immunoglobulin present in normal human serum and the most used format of therapeutic antibodies (Fig. 2). Rituximab against the CD20 receptor is an example of this format of therapeutic antibodies.[12–14]

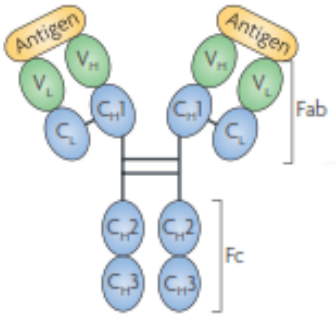


Figure 2 - **IgG structure**. IgG is composed by two heavy and two light chains. Constant regions constitute the Fc domain and variable regions constitute the Fab domain, allowing antigen binding. Adapted from Weiner *et al*, 2010.[34]

1.2.2. Rituximab, the first chimeric IgG anti-CD20

Rituximab is a chimeric murine/human antibody consisting of a glycosylated IgG1 with a human kappa constant regions and murine light and heavy chain variable regions. This antibody is included in the first generation of anti-CD20 antibodies and has three main mechanisms of action: antibody-dependent cell-mediated cytotoxicity (ADCC); complement-dependent cytotoxicity (CDC) and induction of apoptosis.[11,15]

Despite the success of rituximab, various mechanisms of resistance have been developed by the tumoral cells against this antibody, such as the decreasing of CD20 expression on B-cells, structural changes affecting the binding region of the CD20 antibody or alterations in the cell membrane.[11] Thus, new anti-CD20 antibodies have been developed, such as Ofatumumab and Obinutuzumab, which show different immunogenicity, binding to different epitopes on the target with a different affinity (Fig. 3).[16]

These new antibodies discovered for the CD20 receptor can overcome some disadvantages of Rituximab; however they have also a conventional format of IgG and have shown some mechanisms of resistance and immunogenicity.[17] Within this context, there is still an urgent need to identify and develop new antibody

molecules and scaffolds for NHL treatment. A promising alternative to conventional mAbs molecules are recombinant antibody fragments.

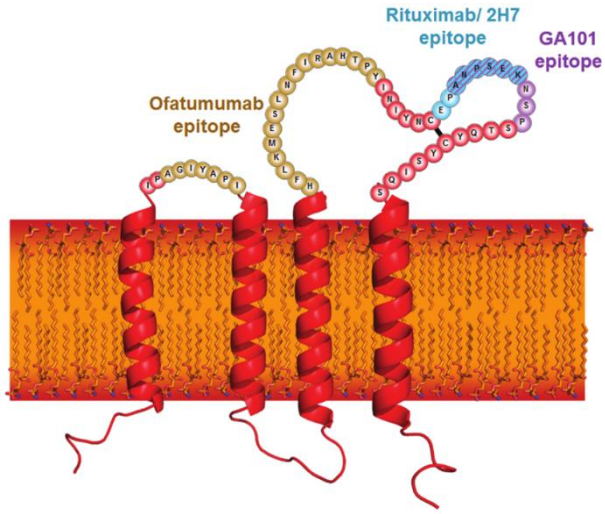


Figure 3 - **Structure of CD20 and epitopes recognized by Rituximab, Ofatumumab and Obinutuzumab** (GA101). Adapted from Klein *et al*, 2013.[15]

1.2.3. Novel therapeutic antibodies

As mentioned above, most marketed antibodies are composed of a full-length IgG molecule that provides for a long half-life and effector functions. However, there is a range of therapeutic applications in which other antibody formats may be more desirable. For instance, in some conditions, a long serum half-life may increase the antibody immunogenicity, and inappropriate activation of Fc-receptor-expressing cells may lead to massive cytokine release and associated toxic effects. In addition, due to their high molecular weight (~150 kDa), IgG antibodies are known to diffuse poorly into tumors and clear slowly from the body. Therefore, to avoid Fc-associated effects in some clinical settings and address the size limitations of IgGs, smaller antibody molecules such as the antigen-binding fragment (Fab), the variable fragment (Fv) and VL or VH single domain antibodies (sdAbs) may be constructed and have become more attractive as therapeutic agents (Fig. 4).[18,19]

1.2.3.1. Single-domain antibodies

Single domain antibodies (sdAbs) are the variable domain of heavy (VH) or light (VL) chains. These fragments are the smallest functional antigen-binding fragments (11kDa to 15kDa) that can be isolated from conventional IgGs.[20] sdAbs show several advantages mainly related with their reduced size. One of them is the easy access to the epitopes on the target surface, which cannot be reached by IgG or larger antibody fragments (ex: Fab or scFv). Furthermore, sdAbs can be expressed efficiently in bacteria as active, soluble and robust proteins. Their reduced size is also an advantage because sdAbs can penetrate solid tissues and tumors more efficiently.[7] Moreover, since sdAbs lack the Fc domain, the non-specific uptake in tissues that highly express Fc receptors is low as well their immunogenicity's. As therapeutic molecules, sdAbs are more stable than full-sized antibodies in circulation, meaning that they could be taken orally or delivered via the pulmonary route. These features allow the administration of higher molar quantities per gram of product, which should provide a significant increase in potency per dose and a reduction in overall manufacturing cost.[21] In addition, sdAbs can be efficiently selected by in vitro screening approaches, such as the Phage Display technology. Taking in consideration all these advantages, sdAbs have become a promising alternative to conventional antibodies for challenging clinical applications such cancer and NHL.

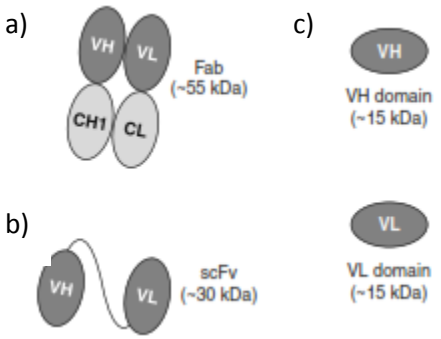


Figure 4 - Structure of antibody domains. a) Antigen- binding fragments (Fabs). b) Single-chain variable fragments (scFv). c) Heavy and light chains (VH and VL domains). Adapted from Aires da Silva *et al,* 2008.[35]

1.3. Selection of monoclonal antibodies

Mouse hybridoma technology, described in 1975 by Kohler and Milstein, enabled the production of monoclonal antibodies through the fusion of myeloma cells with antibody producing B-cells.[22] However,

because hybridoma antibodies are typically originate from murine sources, this limits their therapeutic applications due to human anti-mouse antibody reaction.[23] To overcome this limitation, several strategies have been developed, such as chimerization and humanization.[24] In addition, new methods that use *in vitro* display technology have been introduced, such as antibody phage display and yeast, ribosome and bacterial systems. The phage display has been the most established and is currently the best antibody selection method use in the development of therapeutic antibodies.[25]

1.3.1. Phage display technology

Phage display was initially described by Smith in 1985[26] and is based on genetic engineering of bacteriophages (viruses that infect bacteria) and repeated rounds of antigen-guided selection and phage propagation.[27] In this technique, antibody genes are linked to the amino terminus region of the phage minor coat protein pIII. When expressed, the encoded fusion product is incorporated into the mature phage particle during normal biogenesis. The resulting phage particle expresses antibodies on its surface and contains the antibody encoding gene. The linkage of displayed antibody phenotype with encapsulated genotype via the phage surface, consists in the crucial advantage of this technology.[28]

The most common screening method is based on enriching the phages clones with binding affinity for the target by a process called panning. Each panning consists of multiple rounds of phage binding to the antigen, washing, elution and reamplification of the phage binders. During each round, specific binders are selected out from the pool by washing away non-binders and selectively eluting binding phage clones. Usually, three to six rounds of binding, elution and amplification (panning) are sufficient to generate antibodies with high affinity and specificity (Fig. 5).[27]

Various types of phage display libraries may be used: 1) immune libraries that have been constructed from variable domains of antibody genes of B cells derived from different kinds of immunized animals such as rabbits and mice; 2) naïve libraries, derived from non-immunized donors of B cells, which have been constructed from a pool of V-genes of IgM mRNA;[29] 3) synthetic libraries that also derived from non-immune sources as their ranges were prepared synthetically by combining germ line gene sequences together with randomized complementary determining regions (CDRs) that are responsible for antigen binding.[30] Due to its characteristics, phage display presents itself as a technology with enormous potential for the selection of promising antibodies.

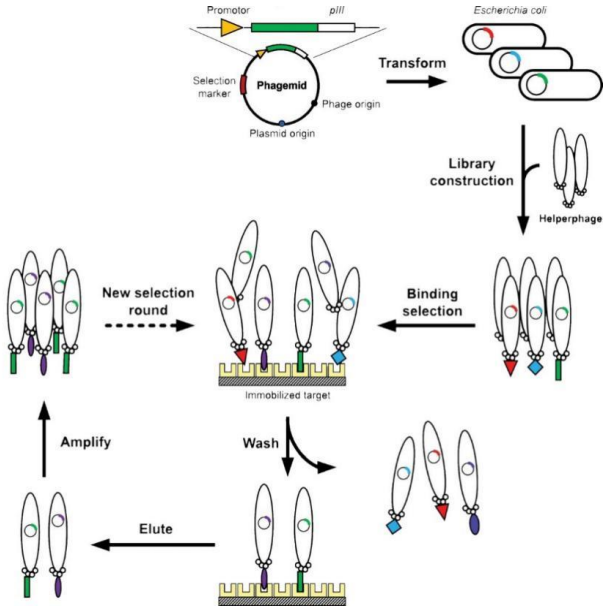


Figure 5 – Schematic representation of a round of selection (panning) for phage display. Adapted from Ruigrok *et al*, 2011.[36]

1.4. Objectives

Due to the disadvantages that antibodies, currently in the market for the CD20 receptor in B-cell malignancies possess, such as size, immunogenicity, as well as the resurgence of several resistances, the development of more promising antibodies for this target is an urgent need. Taking this into account, the aim of this project was the development and characterization of a new antibody against CD20 that could be efficiently used in the future in the treatment of NHL and other B-cell malignancies. To achieve this goal, the potential of rabbit derived sdAbs for therapeutic applications was explored and a subtractive cell phage display screening was used as a functional selection. Briefly, the project consisted in following tasks/aims:

1) Rabbit immunization with CD20;
2) Construction of immune sdAb libraries against the CD20;
3) Selection of sdAbs by phage display technology;
4) Screening and characterization of selected sdAbs against CD20.

The present work was realized in Faculty of Veterinary Medicine of University of Lisbon, Faculty of Sciences of University of Lisbon and Technophage SA, headquartered in Instituto de Medicina Molecular (IMM), Faculty of Medicine of University of Lisbon.

2. Material and Methods

2.1. Cell culture

Cell lines used in this work were cultured with RPMI 1640 (Lonza) or DMEM (Lonza) supplemented with 10% fetal bovine serum, penicillin and streptomycin. Cells were maintained in a humidified incubator at 37°C and 5% CO_2.

Cells were lysed with Ripa buffer (25mM TrisHCL pH 7.6, 150mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS) (Thermo Scientific) to obtain the total cell extract. Total protein was quantified using the NanoDrop™ 2000c Spectrophotometer (Thermo Scientific).

2.2. Transfection of human CD20 into HEK 293T cells

The HEK 293T cells (ATCC) were transfected using the FuGENE® HD (Roche), according to the manufacturer's instructions. For that purpose, 8 µl of FuGENE® HD was added to the 80 µl of medium and 3 µg of pFUGW vector CD20, to a final volume of 100 µl. Then, the mixture was added to each well containing the cells plated. The plate was swirled and the cells incubated for 4 hours. After incubation, the medium was changed and the cells incubated for 48 hours. After 48 hours, cells were washed and resuspended in PBS 1x and injected in the rabbit.

2.3. Rabbit immunizations

A New Zealand White rabbit was immunized with the CD20 peptide (141-184 aa) and HEK 293T cells transfected with CD20 vector through subcutaneous injections. The injections were administrated in 2 - 3 weeks for 3 months.

To evaluate the serum specificity and titer, serum collected before and after bleeding from the rabbits was analyzed by ELISA using either Raji cells (ATCC) or the CD20 peptide as antigen, and Peroxidase-conjugated goat anti-rabbit antibody (Jackson Immune Research) as secondary antibody.

Serum was purified using an AKTA^{tm} start system in order to evaluate serum activity and titer against the CD20 receptor. For protein purification, 1 ml HiTrap™ Protein A HP was used. The column was equilibrated with PBS 1x, pH 7.4. Bound proteins were eluted from the column with 0.1 M citric acid. To the collecting tubes, 200 µl of 1M tris-HCl, pH 9.0, were added to neutralize samples.

2.4. sdAbs immune library construction

2.4.1. Isolation of Total RNA

Five days after the final boost, spleen (SP) and bone marrow (BM) were harvested, homogenized and incubated in 10 ml of Tri reagent at room temperature for 5 minutes. Then, 20 ml of Tri reagent was added to each sample, and then centrifuged at 2500 g for 10 minutes at 4°C. The pellets were discarded, and for each supernatant, 3 ml of 1-Bromo-3-chloropropane BCP was added. Samples were agitated and incubated for 15 minutes at RT. Next, they were centrifuged at 17500 g for 15 minutes at 4°C. The colorless aqueous phase of each sample was transferred to a new centrifuge tube containing 15 ml of isopropanol. Samples were agitated and incubated for 10 minutes at RT. The samples were then centrifuged at 17500 g for 10 minutes at 4°C. The supernatant was discarded and 30 ml of 75% ethanol was added to the pellet. Centrifugation at 17500 g for 10 minutes, at 4°C, was then performed. Supernatants were then discarded, and pellets air-dried at RT. Finally, each pellet was resuspended in 500 µl of RNase-free water. The RNA concentration and purity was evaluated by measuring the samples absorbance at 260 nm in a NanoDrop™ 2000c Spectrophotometer (Thermo Scientific) and on an agarose gel electrophoresis, respectively.

2.4.2. First strand cDNA synthesis

Complementary DNA (cDNA) was synthetized by reverse transcription from total RNA using the Transcriptor High Fidelity cDNA Synthesis Kit (Roche), according to the manufacturer's instructions. After synthesis, cDNA was quantified using the NanoDrop™ 2000c Spectrophotometer (Thermo Scientific) through the measurement of the samples absorbance. The integrity of the synthetized cDNA was evaluated on a 1% agarose gel in TAE 1% by ethidium bromide staining under UV light with Chemidoc™ XRS+ (Bio-Rad).

2.4.3. PCR amplification of single domains antibody genes

PCR was performed to amplify the heavy-chain variable (VH) and light-chain variable regions (VL) from the rabbit's cDNA. The following conditions were used: an initial denaturation step, at 98°C for 30 seconds; a denaturation for 10 seconds at 98 °C; an annealing for 30 seconds at 56 °C and an extension for 60 seconds at 72°C, for 30 cycles; final extension for 10 minutes at 72°C. For the amplification of the rabbit's VH coding sequences, a four primer combination was used and for the VL, a ten primer combination (shown in the supplementary data section and Fig. 6). For each reaction, 500 ng of cDNA was used in a mix with 0.5 µM of Primer Forward and Reverse (20 pmol/µl), buffer 1x (5X), 3 mM of $MgCl_2$ (50 mM), 200 µM of dNTPs (10 mM), 0.02 U/µl of Phusion High Fidelity DNA polymerase (Thermo Scientific) and miliQ water to a final volume of 50 µl.

Each reaction was evaluated on a 2% agarose gel by ethidium bromide staining under UV light with Chemidoc™ XRS+ (Bio-Rad).

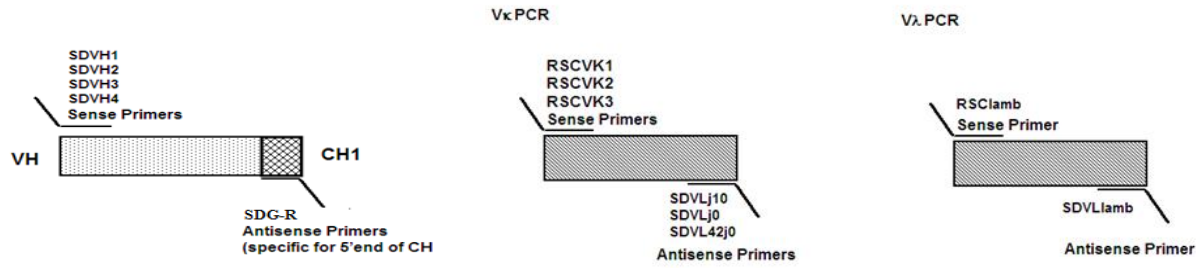


Figure 6 – Generation of VH and VL antibody fragments by PCR.

2.4.4. Precipitation of PCR products

To concentrate the PCR products, DNA precipitation with Sodium Acetate was performed. A mixture with 1/10 of the volume of 3 M NaOAc, 2.5 volumes of cold absolute EtOH and 1 µl of glycogen (Roche) was added to the DNA, and then incubated overnight at -80°C. After incubation, the mixture was centrifuged at 13200 rpm for 1 hour at 4°C, the supernatant was discarded, 1 ml of cold 70% EtOH was added to the pellet for washing and then centrifuged one more time. The supernatant was removed and the pellet was dried and resuspended in miliQ water (the amount was adjusted to each library).

2.4.5. Purification of PCR products

In order to purify DNA, PCR products were separated by electrophoresis on a 2% low-melt agarose gel in TAE 1x buffer. DNA with the expected size was extracted from the gel with a blade scapel and purified with the QIAquick® Gel Extraction Kit (Qiagen), according to the manufacturer's instructions. Then, DNA was precipitated as described above.

DNA concentration was determined using NanoDrop™ 2000c Spectrophotometer (Thermo Scientific) through the measurement of the samples absorbance.

2.4.6. Digestion of vector DNA and PCR products with restriction enzyme *Sfi*I

For the PCR products digestion, 10 µg of purified DNA, 160 units of Restriction Endonuclease *Sfi* I (40 U/µl) (Roche), 20 µl of buffer M (10x) and miliQ water were used to obtain a final volume of 100 µl. For the pComb3x vector digestion, 10 µg of vector, 120 units of *Sfi* I, 20 µl of buffer M (10x) and miliQ water were used to obtain a final volume of 100 µl. Both digestions were incubated overnight at 50°C.

After digestion, vector and PCR products were purified on a 0.8% and 2% low melt agarose gel, respectively. DNA with the expected size was then removed from the gel and extracted by QIAquick® Gel Extraction Kit (Qiagen), and precipitation once again performed. Later, digested products were quantified using the NanoDrop™ 2000c Spectrophotometer (Thermo Scientific).

2.5. Phage display library

2.5.1. Cloning of PCR products into pComb3x vector

In order to clone VL and VH PCR libraries into pComb3x, a ligation was performed using 1.4 µg of vector, 4.2 µg of insert, 20 µl of ligase buffer (10x), 5 µl of T4 DNA ligase (BioLabs) and miliQ water to obtain a final volume of 200 µl. The ligation was incubated overnight at 16°C. The following day, DNA precipitation with sodium acetate was performed, following the steps described above, and DNA pellet was resuspended in 20 µl of miliQ water.

Each 5 µl of ligation was inserted into 60 µl of electrocompetent *E.coli* ER2738 (Lucigen®) by electroporation. For this, an electrical shock at 1.8 kV, 25 µF and 200 Ω was performed. After, 1 ml of SOC medium was added to each cuvette and the final volume was adjusted to 5 ml. The reaction was incubated for 1 hour at 37°C, agitated at 200 rpm. After 1 hour, 4.5 µl of ampicillin (100 mg/ml) was added and incubated at 37°C, agitated at 200 rpm. Next, 10 ml of pre-warmed SB medium with 3 µl of ampicillin (100 mg/ml) and 15 µl of tetracycline (5 mg/ml) were added to the culture and incubated for 1 hour at 37°C, agitated at 200 rpm. After incubating for 1 hour, 2 ml of helper phage VCSM13 and 183 ml of pre-warmed SB medium with 92.5 µl of ampicillin (100 mg/ml) and 185 µl of tetracycline (5 mg/ml) were added to the culture. The culture was agitated at 200 rpm for 1.5 hours at 37°C. Finally, 280 µl of kanamycin was added and incubated overnight at 37°C, agitated at 200 rpm.

The following day, the phages produced were recovered by precipitation using 8 g of PEG-8000 and 6 g of NaCl for 1 hour on ice. Next, the phages were centrifuged at 8000 rpm for 20 minutes at 4°C. The supernatant was discarded and the pellet dried by inverting on a paper towel, resuspended in 2 ml of PBS 1x and transferred to a 2 ml eppendorf. To eliminate the remaining bacteria, a final centrifugation was performed at 13200 rpm for 5 minutes. Lastly, the phages supernatant was filtered through a 0.22 µm filter.

2.5.2. Confirmation of library insert efficiency and diversity

To confirm the library insert efficiency and diversity, PCR was performed. Individual colonies were picked into PCR tubes containing 20 µl of H_2O and incubated for 5 minutes at 94°C in a thermocycler. Then, PCR was performed using 5 µl of colony sample, 0.8 µl of primer RSC-F (20 pmol), 0.8 µl of primer RSC-B (20 pmol), 5 µl PCR Buffer (5X), 1.5 µl of MgCl2 (50 mM), 0.5 µl of dNTPs (10 mM), 0.2 µl of Phusion High Fidelity DNA polymerase (Thermo Scientific) and miliQ water to a final volume of 25 µl. The amplification conditions used are mentioned above. PCR products were evaluated on a 2% agarose gel by ethidium bromide staining under UV light with Chemidoc™ XRS+ (Bio-Rad).

2.6. Selection of specific sdAbs against the CD20 receptor by phage display

For each VH and VL library, four selection rounds were performed and two types of cells were used for selection: positive cells - Raji; negative cells - HEK 293T.

A total of 5 x 10^6 cells were used for each type. Cells were washed twice with PBS 1x and blocked with PBS/BSA 1%, at RT for 15 minutes, as well as the fresh phages for each library. Next, 300 µl of blocked phages were added to the blocked HEK 293T cells, and incubated for 15 minutes at 37°C. After incubation, the cells were centrifuged at 300 g for 5 minutes at RT, the supernatant recovered and added to the Raji cells. Raji cells were incubated with phages for 1.5 hours at 37°C, and the tubes inverted every 30 minutes to mix. After incubation, the tubes were centrifuged at 300 g for 5 minutes, and the cells pellet washed with PBS 1x, followed by several centrifugations. In the first panning, three wash steps were performed. In the second panning, six wash steps were performed, and in the third and fourth pannings, the cells were washed eight times. After the final wash, the cells were resuspended in 200 µl of pre-warmed trypsin and incubated at 37°C for 7 minutes. Next, the trypsin was immediately diluted with 1.2 ml of PBS 1x and the cells were centrifuged at 300 g for 5 minutes. The phages in the supernatant were named trypsin phages (or binder phages) and stored at 4°C. Then, the cells pellet was washed twice with PBS 1x and the supernatant discarded. To recover internalized phages, a cell lysis with 200 µl of PBS-Tween20 0.5% was performed, and the cells were incubated for 7 minutes at 37°C. Next, the cells were centrifuged at 300 g for 5 minutes and the supernatant containing internalized phages recovered and stored at 4°C.

After a round of panning, a phage reamplification was performed. To reamplify, it was necessary to grow *E.coli* ER2738 bacteria from a fresh plate in 10 ml of SB medium containing 20 µl of tetracycline at 37°C at 200 rpm until reaching an O.D. of approximately 0.6 - 0.7. Then, each phage (trypsin or internalized) was infected with 4 ml of bacteria and incubated for 15 minutes at 37°C. Next, 4 ml of pre-warmed SB medium with 1.6 µl of ampicillin and 8 µl of tetracycline was added to the culture. The culture was incubated for 1 hour at 37°C, agitated at 200 rpm. Then, 2.4 µl of ampicillin was added to the culture and incubated in the same conditions for 1 hour. After incubation, 1 ml of VCSM13 helper phage was added to the culture and incubated for 15 minutes at 37°C. Then, the culture was transferred to a 250 ml Erlenmeyer containing 91 mL of pre-warmed SB medium with 46 µl of ampicillin and 184 µl of tetracycline. The culture was incubated for 1.5 hours at 37°C, agitated at 200 rpm. After, 140 µl of kanamycin was added and the culture incubated overnight, agitated at 200 rpm at 37°C.

The following day, phage precipitation was performed as described previously, and then a new round of selection was performed (Fig. 7).

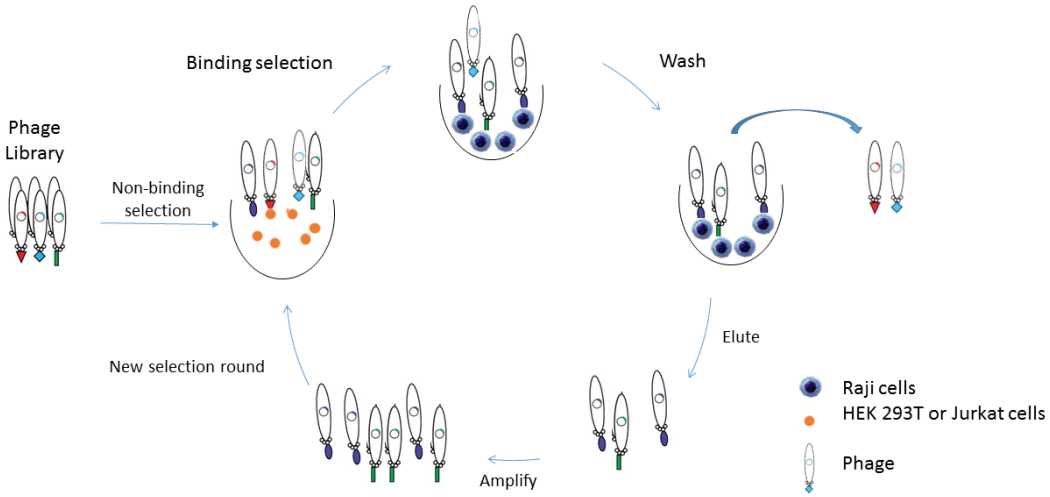


Figure 7 - Schematic representation of a round of selection (panning) for phage display.

2.7. Screening and characterization of sdAbs against the CD20 receptor

2.7.1. Cloning of selected sdAbs genes against the CD20 receptor into the pT7-PL (pT7-peptide leader) vector

DNA into pComb3x was extracted from bacteria of the fourth panning of each phage display screening using a QIAprep® Spin Miniprep Kit (Qiagen), according to the manufacturer's instructions, as described above, and then a digestion with restriction enzyme *Sfi* I was performed, according to the conditions also described above. After, the digested products were purified in order to separate the vector from the fragments with the desired size. Next, the purified DNA was inserted into pT7-PL vector using the T4 ligase enzyme, 50 ng of vector, 150 ng of VL or VH digested products, according to the conditions mentioned above.

Then, transformation was performed, using 2 µl of ligation and 100 µl of *E.coli* BL21 cells. The conditions used were the same as the ones described previously. After incubation for 1 hour at 37°C, 200 rpm with SOC medium, 10 and 100 µl of each ligation were plated in LB + Amp plates. The next day, a confirmation of library insert efficiency was performed, as mentioned above.

2.7.2. Expression and selection of specific antibodies

The colonies obtained previously through transformation into *E.coli* BL21 cells were picked and incubated overnight at 30°C on 100 µl of SB medium containing Overnight Express™ Autoinduction System 1 (Novagen®) and 100 µg/ml of Ampicilin. Next day, 40 µl of BugBuster (Roche) containing anti-protease cocktail-EDTA free inhibitors (Roche) were added and incubated for 30 minutes at 4°C. Next, the plates were centrifuged at 2000 rpm, for 10 minutes, and the supernatant was used to perform an ELISA assay. Three types of parameters were analyzed: binding of antibodies to the antigen, level of expression and unspecific binding. To evaluate the binding to the antigen (CD20), plates with the total Raji cell extract attached were used, then washed three times with PBS 1x and blocked with BSA 3% for 1 hour at 30°C. Then, the plates were washed again five times and the supernatant containing VLs or VHs was incubated for 1 hour at 30°C. The plates were washed again 5 times and anti-HA-HRP antibody (Roche) was added.

After 1 hour of incubation, the plates were washed 5 times, ABTS (Roche) substrate solution was added and the optical density at 405 nm was measured at different time points. To evaluate the level of expression, the same process was carried out except the coating of the plates with antigen. To evaluate the unspecific binding, in turn of the antigen, BSA 3% was added. Rabbit serum and anti-CD20 antibody were used as positive controls. BL21 extracts and SB medium were used as negative controls.

The best clones were chosen and a new ELISA assay with extracts and Raji cells were performed to evaluate the binding of antibodies to the CD20 in a native conformation.

2.7.3. Sequencing of anti-CD20 clones

Sequencing was performed by GATC using the pComb3x ATG primer. To translate to amino acid sequences and, in order to evaluate their homology, the "Vector NTI" software was utilized.

2.7.4. Confirmation of the specificity for the CD20 by Western Blot

In order to verify if the selected clones were specific to the CD20 receptor, Western Blot was performed. Membranes blotted with the Raji, HEK 293T and Jurkat cells were incubated with the clones using anti-HA-HRP (Roche) as secondary antibody.

The clones for the Western Blot were purified using a Talon® spin column. The column was equilibrated with the equilibration/wash buffer, containing 50 mM of NaH_2PO_4, 300 mM of NaCl and 20 mM of Imidazole. Bound proteins were eluted from the column with 50 mM of NaH_2PO_4, 300 mM of NaCl and 150 mM of Imidazole.

3. Results and discussion

3.1. Rabbit immunizations

The rabbit antibody repertoire has been used for decades in diagnostic applications in the form of polyclonal antibodies. Now it also holds great promise as a source for generation of therapeutic mAbs. Moreover, over the past decade, rabbit derived sdAbs have been shown to have a great potential for therapeutic applications.[31,32] Therefore, in this project we have explored rabbit derived sdAbs to develop new and potent antibodies against the CD20 receptor. In order achieve the proposal goal, one rabbit was immunized with the CD20 peptide and HEK 293T cells transfected with the CD20 vector. The rabbit was immunized for 84 days and euthanasia performed on day 89 according with the schedule shown on Table 1.

Table 1 - Rabbit immunization schedule.

Day	Rabbit immunization	Action
0	26th February 2015	Pre-Bleed + Immunization with the CD20 peptide
14	12th March 2015	Bleed 1 + Immunization with 293T cells transfected with the CD20 vector
28	26th March 2015	Bleed 2 +Immunization with 293T cells transfected with the CD20 vector
42	9th April 2015	Bleed 3
56	23rd April 2015	Bleed 4 +Immunization with 293T cells transfected with the CD20 vector
70	7th May 2015	Bleed 5 +Immunization with 293T cells transfected with the CD20 vector
77	14th May 2015	Bleed 6
84	21st May 2015	Final Immunization with 293T cells transfected with the CD20 vector
88-89	25th May 2015	Terminal Bleed+ Euthanasia

To evaluate serum specificity and titer, serum was collected according to the table 1 and analyzed by ELISA. The results corresponding to ELISA with pre and final bleed against the Raji cells are shown in Figure 8. Raji cells are B cells that express the CD20 receptor. These results show that the final bleed serum recognized the Raji cells contrarily to pre-bleed that did not, which also occurs in the ELISA with the CD20 peptide (Fig. 8). It also shows that immunizations resulted in a strong immune response as the serum titer was higher, as was expected for rabbit immunizations with cells.[27] Moreover, it is also import to mention that no bind was observed for HEK 293T cells (data not shown). Therefore, the rabbit was euthanized and BM and SP extracted.

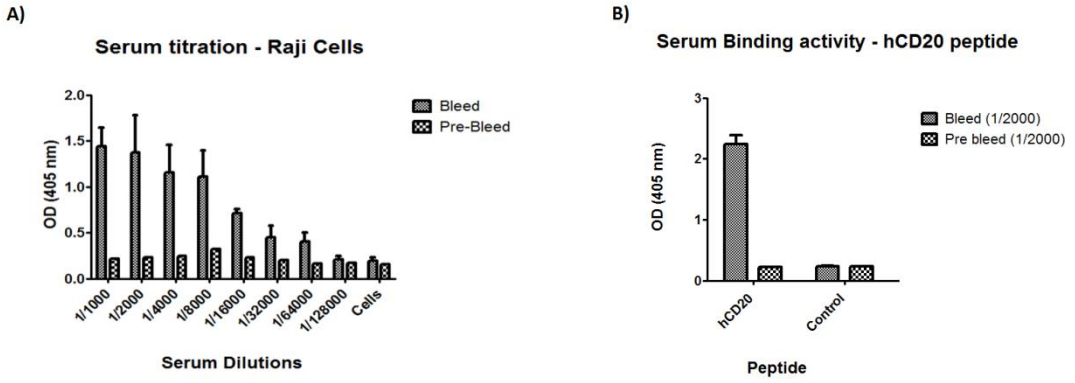


Figure 8 - **Titration and binding activity of serum antibodies corresponding to pre and final bleed.** Antisera from the immunized rabbit was analyzed for binding to Raji cells (A) and human CD20 peptide (B) by ELISA using HRP-conjugated goat anti-rabbit Fc polyclonal antibody as secondary antibody. Data were obtained by Abs measurement at 405 nm with reference a 492 nm. Values represent mean±standard deviation of triplicates (n=3) for each condition, and similar results were obtained in two independent experiments.

3.2. sdAbs immune library construction

3.2.1. RNA extraction and cDNA synthesis

Immunized sdAbs libraries were constructed using the RNA and cDNA obtained from the rabbit BM and SP. These sources were chosen because they have been shown to be a major repository of plasma cells that secretes antibodies used for construction of immune libraries.[27] RNA and cDNA were extracted and synthetized as described in the material and methods section. In order to evaluate their concentration and purity, absorbance at 260 nm was measured, and integrity analyzed in agarose gel electrophoresis. RNA and cDNA yields and purity were as expected and according to data obtained for rabbit immunizations (Table 2 and Fig. 9).[27]

Table 2 - Quantification of RNA and cDNA from BM and SP of immunized rabbit.

	RNA [ng/µl]	cDNA [ng/µl]
Bone marrow	563,9 (Total: 2 mg)	1485,1 (Total: 44.5 µg)
Spleen	799 (Total: 4 mg)	2529,3 (Total: 75.8 µg)

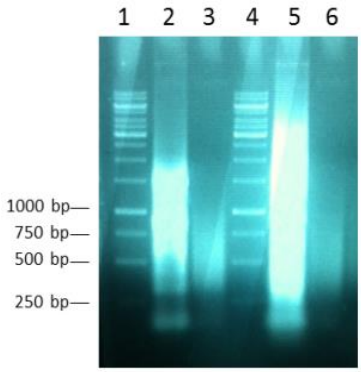


Figure 9 - Samples of RNA extracted and cDNA synthetized from the BM and SP were eletrophoretically separated on a 1% agarose gel. (1 and 4 - Ladder 1kb; 2 - RNA BM; 3 - cDNA BM; 5 - RNA SP; 6 - cDNA SP)

3.2.2. PCR amplification of single domains genes

Immune sdAbs libraries were constructed using knowledge of the immunoglobulin variable regions sequence.[27] Light (VL) and heavy (VH) chains were amplified separately according to the conditions described in the material and methods section. The VL regions were amplified from BM and SP cDNA using the VL sense and reverse primers generating fragments with 350 bp. The VH regions were amplified also from BM and SP cDNA by using the VH sense and reverse primers, originating fragments with 400 bp. PCR products had the expected size (Fig. 10).[27] The amplification of the VH and VL families contributed to a higher diversity of immune libraries. Therefore, several PCRs were performed to achieve such diversity. All PCR products had a symmetric *Sfi* I sites on the 5' and 3' ends, which was required for cloning into the pComb3x vector.

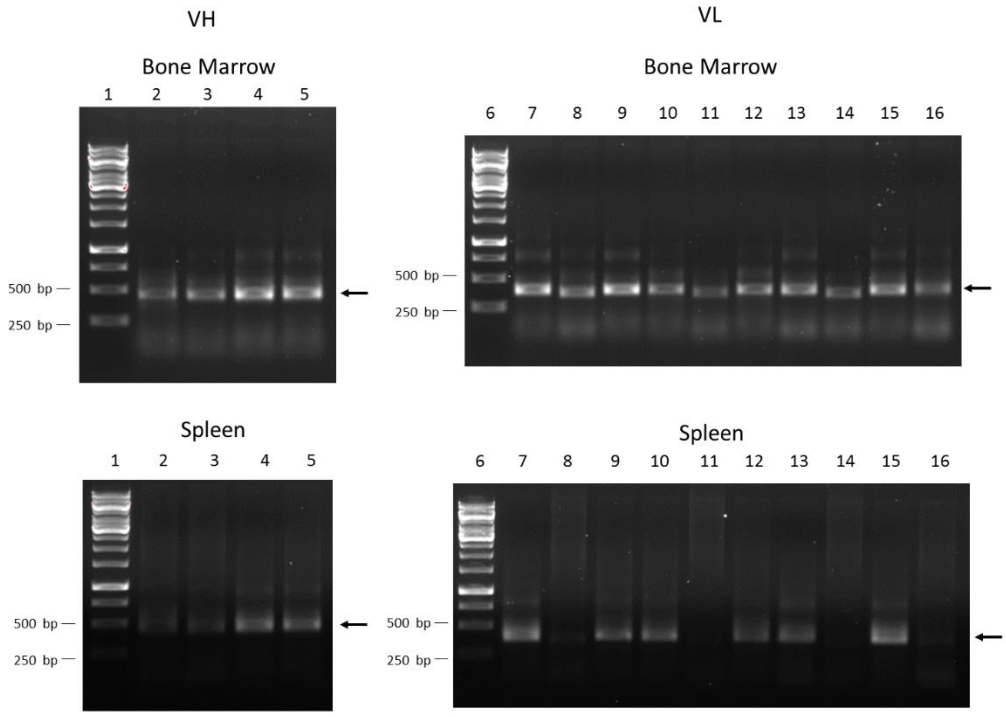


Figure 10 - PCR amplification of VL and VH families from BM and SP cDNA. Samples were run and separated on a 2% agarose gel. Amplification of VLs generated fragments with approximately 350 bp and VHs with about 400 bp. (1 and 6 - Ladder 1 kb; 2 - SDVH1+SDG; 3 - SDVH2+SDG; 4 - SDVH3+SDG; 5 - SDVH4+SDG; 7 - RSCVK1+SDVLj0; 8 - RSCVK1+SDVLj10; 9 - RSCVK1+SDVL42j0; 10 - RSCVK2+SDVLj0; 11 - RSCVK2+SDVLj10; 12 - RSCVK2+SDVL42j0; 13 - RSCVK3+SDVLj0; 14 - RSCVK3+SDVLj10; 15 - RSCVK3+SDVL42j0; 16 - RSCλ+ SDVLλ). The arrows indicate the VL or VH PCR products.

3.2.3. Purification and cloning of PCR products into pComb3x vector

After amplification, PCR products were purified in agarose gel as previously described. 45.2 µg and 25.7 µg of DNA were recovered from the purification of VL BM and SP, respectively. Regarding the purification of VH BM and SP, 20.5 µg and 16.2 µg were recovered, respectively (Fig. 11).

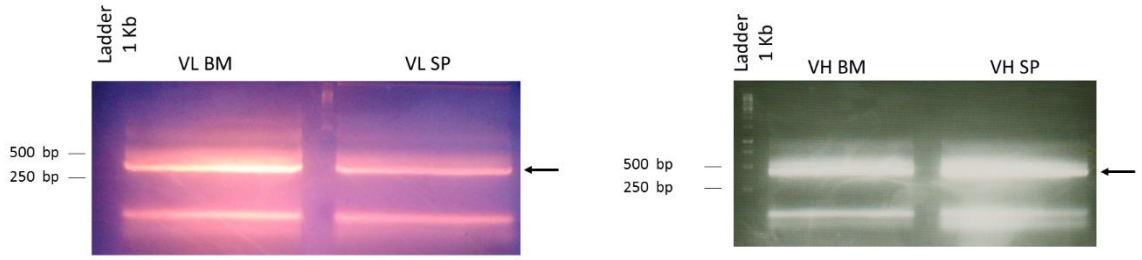


Figure 11 - Purification of VL and VH domains after amplification in 2% low melting point agarose gel. For the VLs, the fragment corresponding to 350 bp was extracted from the gel. Regarding VH, the fragment corresponding to 400 bp was extracted from the gel. The arrows indicate the VL or VH products purified. Ladder 1kb as the DNA size control.

The purified DNA´s were then digested with the *Sfi* I restriction enzyme and purified again. 4.16 µg and 4.5 µg of DNA were recovered from the purification of VL BM and VL SP, respectively. Regarding the purification of VH BM and SP, 5.3 µg of DNA were recovered from each one. The agarose gels regarding the purifications are presented below (Fig. 12).

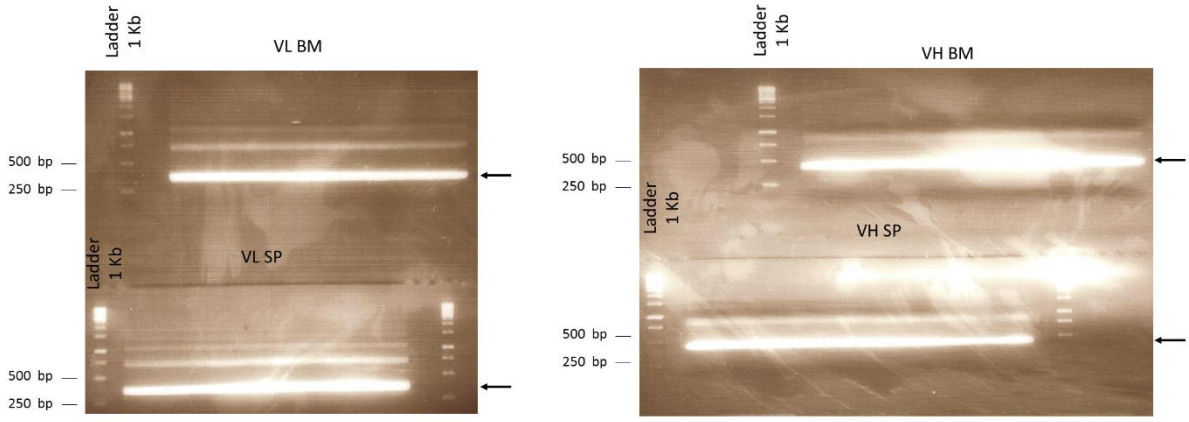


Figure 12 - **Purification of VL and VH domains after Sfi I restriction in 2% low melting point agarose gel.** For the VLs, the fragment corresponding to 350 bp was extracted from the gel. Regarding VH, the fragment corresponding to 400 bp was extracted from the gel. The arrows indicate the VL and VH domains. Ladder 1kb as the DNA size control.

After digestion and purification, the DNA recovered was ligated into the pComb3x. Ligation was then inserted into *E.coli* ER2738 cells by electroporation, which is the most efficient method for transforming *E.coli* with plasmid DNA.

The number of independent transformants is very important, as it is used to describe the libraries' complexity, allowing to determine the quality of the library. For the libraries derived from immune animals, the number of individual transformants must be in the range of 10^7 to 10^8.[27] For the VL BM and SP libraries, the diversity was 8 x 10^7 and 2 x 10^7 transformants, respectively. The diversity obtained for the VH BM and SP libraries were 1.3 x 10^7 and 3.8 x 10^6 transformants. To confirm the libraries insert efficiency, PCR was performed as described in the material and methods section, using 10 clones of each library. An efficiency of 100% was obtained, meaning that all transformants had the insert cloned (Fig. 13). Because both libraries had a number of approximately 10^7 transformants, we can conclude that both libraries were efficiently constructed and are also representative, as they are highly diverse and complex.

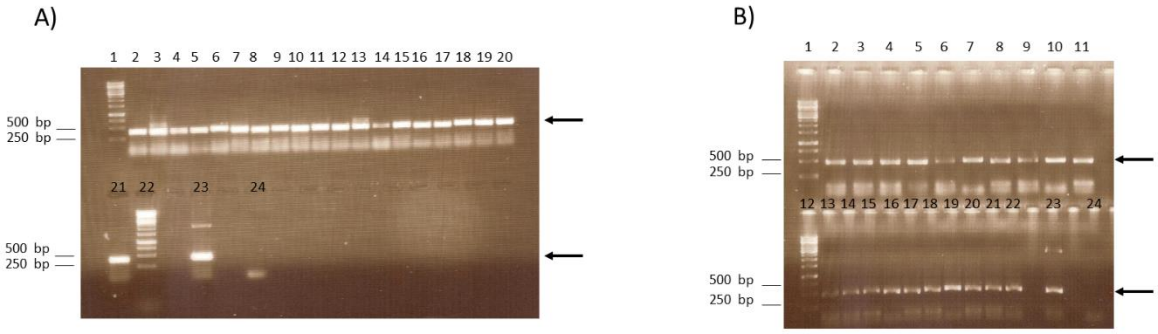


Figure 13 - Confirmation of the libraries insert efficiency for VL BM and SP libraries (A) and for VH BM and SP libraries (B). Samples were run and separated on a 1% agarose gel. The arrows indicate the VL and VH products. A) 1 and 22- Ladder 1 kb; 2 to 11- VL BM clones; 12 to 21- VL SP clones; 23- C^+ ; 24- C . B) 1 and 12- Ladder 1 kb; 2 to 11- VH BM clones; 13 to 22- VH SP clones;

3.3. Selection of specific sdAbs by phage display

After finishing the libraries construction, selection of specific sdAbs for the CD20 receptor were performed by a subtractive cell phage display as described in the material and methods section. This cell phage display screening was based on the previously Carlos Barbas studies [27] were they reported a novel whole-cell selection protocol with negative and positive selection steps (subtractive phage display) designed to remove antibodies reacting with common antigens. The phage display conditions implemented in our project are summarized in Table 3. Briefly, the process was performed separately for the VL and VH libraries. The BM and SP libraries of each one were mixed together once they had a similar diversity. Four pannings were then performed with a subtractive selection with cells that do not express the CD20. For that, three types of cells (Raji, HEK 293T and Jurkat cells) were used that express different

receptors on their surface, contributing to a better selection of antibodies. The Raji cells were used in the positive selection since they are a stable B cell line that expresses the CD20 receptor. The HEK 293T and Jurkat cells do not express the CD20 receptor, so they were used in the subtractive selection for elimination of non-specific antibodies that bind to other receptors. The increase in the number of washes during the pannings also contributed to the removal of the antibodies that had a weak ligation to the cells, leading to the selection of the best antibodies. The elution's performed were done to select for binders and internalized antibodies. Table 3 and Figure 14 show the results obtained for the VLs and VHs phage display. As shown, in all pannings there were a lower number of phages in the output titers compared to the input (~$10^{11\text{-}12}$ to ~$10^{5\text{-}4}$ pfu). Moreover, it was also observed a decrease in the output phages from the 1st panning to the 2nd panning. Then, from the 2nd to the 3rd panning, the number of obtained phages maintained, and only increases in the 4th panning. This profile is expected for a phage display and demonstrated that the conditions implemented were leading to enrichment for specific antibodies-phages to our target.

Table 3 - Results obtained from VL and VH domains selection by phage display and conditions used in each panning.

		1st Panning	2nd Panning	3rd Panning	4th Panning
Input (Phages /ml)	VH	$4,2x10^{11}$	Binders- $2,6x10^{11}$ Internalized- $1,1x10^{11}$	Binders- $1x10^{12}$ Internalized- $2,9x10^{11}$	Binders- $1x10^{12}$ Internalized- $1,9x10^{11}$
	VL	$2x10^{12}$	Binders- $2x10^{11}$ Internalized- $2x10^{12}$	Binders- $2,1x10^{11}$ Internalized- $1x10^{12}$	Binders- $1,6x10^{11}$ Internalized- $7x10^{10}$
Output (Phages/ml)	VH	Binders- $8,2x10^{5}$ Internalized- $4,5x10^{4}$	Binders- $6,7x10^{4}$ Internalized- $1x10^{4}$	Binders- $5,2x10^{4}$ Internalized- $1x10^{4}$	Binders- $4,2x10^{5}$ Internalized- $3,6x10^{4}$
	VL	Binders- $2x10^{5}$ Internalized- $3x10^{4}$	Binders- $1,5x10^{4}$ Internalized- $< 10^{4}$	Binders- $5,2x10^{4}$ Internalized- $< 10^{4}$	Binders- $1,6x10^{5}$ Internalized- $1x10^{4}$
Conditions		HEK 293T- 15 min Raji – 1:30 h Wash: 3x PBS Trypsin- 7 min PBS/Tween – 7 min	HEK 293T- 15 min Raji – 1:30 h Wash: 6x PBS Trypsin- 7 min PBS/Tween – 7 min	HEK 293T- 15 min Raji – 1:30 h Wash: 8x PBS Trypsin- 7 min PBS/Tween – 7 min	HEK 293T- 15 min Jurkat- 15 min Raji – 1:30 h Wash: 8x PBS Trypsin- 7 min PBS/Tween – 7 min

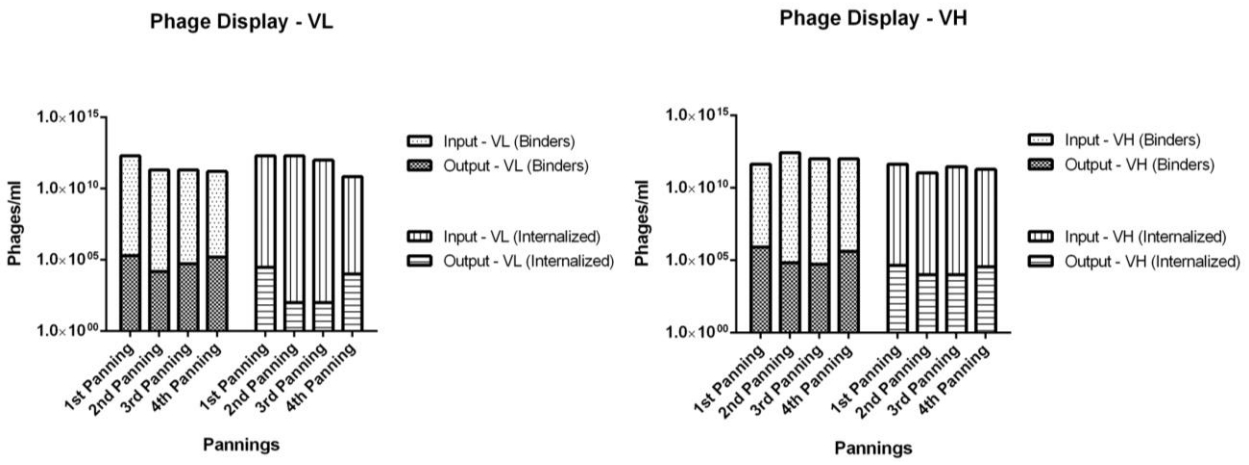


Figure 14 – Graphs illustrating the results obtained from VL and VH domains selection by phage display.

In order to confirm if the obtained sdAbs expressed on the phages surface were specific for the CD20 receptor, western blot with amplified sdAbs expressed on the phages surface from the last panning was performed (Fig. 15). Membrane blotted with the protein extract of the Raji cells was incubated with the sdAbs-phages that were used as primary antibody, and as secondary the anti-M13-HRP (Roche). The results for the VL binders phages are presented below and show that the VL-phages recognized proteins in the cell extract in the range of 35 to 48 kDa.[8] The CD20 protein has a molecular mass of approximately 35 kDa, so it seems that the VL-phages are recognizing the CD20 receptor. Regarding the VH phages, western blot had no success and no bands were obtained (data not shown).

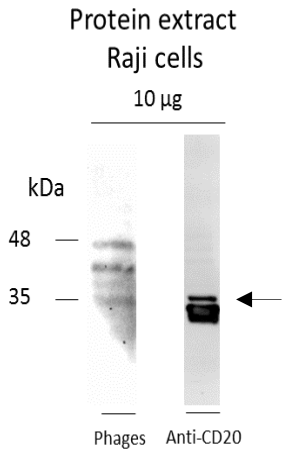


Figure 15 - Western blot analysis of the total cell extract from the Raji cells with the selected VL phage to evaluate the phage binding to the proteins of the cells. The phages were detected using the anti-M13 antibody. The arrow indicates the band corresponding to the CD20 receptor.

3.4. Screening and characterization of sdAbs against the CD20 receptor

3.4.1. Cloning of the selected sdAbs genes against the CD20 receptor to the pT7-PL (pT7-peptide leader) vector

After selection by phage display, it was necessary to select among the pool of antibodies against the CD20, those that had a better expression and binding to the target. To be able to select, it was necessary to clone the DNA corresponding to the VL and VH domains from the last panning into the pT7-PL vector. This vector has a peptide leader, allowing the expression of the sdAbs to be directed to the periplasmic space, making it possible to evaluate the expression level of sdAbs. DNA was digested with the restriction enzyme *Sfi* I, and purified on 2% agarose gel. The results are demonstrated in Figure 16.

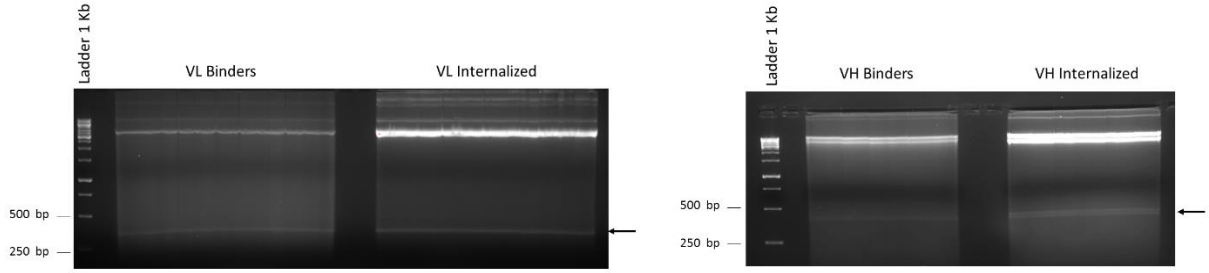


Figure 16 - **Purification of the VL and VH domains selected by phage display after Sfi I digestion.** For the VL, the fragment correspondent to the 350 bp was recovered from the gel, and for the VH, the correspondent to the 400 bp. Each sample was separated on a 2% low melting point agarose gel. The arrows indicate the VL or VH domains selected. Ladder 1kb as the DNA size control.

After purification, DNA was ligated into pT7-PL vector and inserted into *E.coli* BL21 cells by electroporation. In order to evaluate cloning efficiency, PCR was performed. An efficiency of 100% was obtained, meaning that all clones had a cloned insert (Fig. 17).

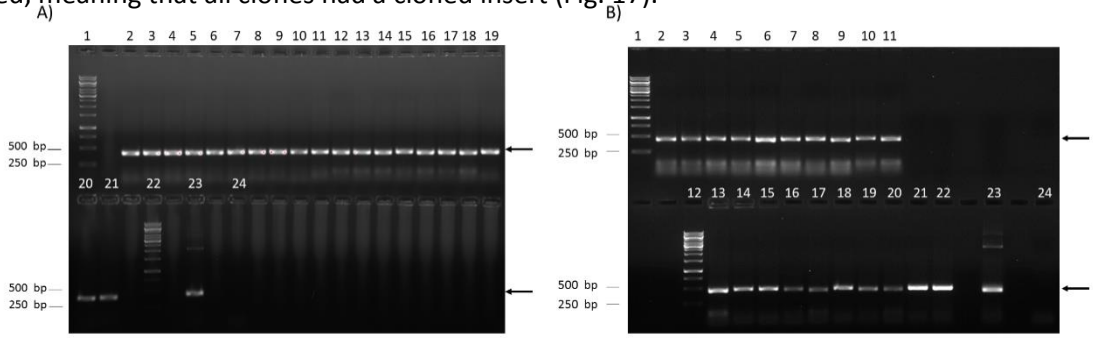


Figure 17 - **Confirmation of the percentage of inserts in VL (A) and VH (B) clones obtained after cloning into pT7-PL vector and transformation into E.coli BL21 cells.** Each clone was eletrophoretically separated on a 1% agarose gel. The arrows indicate the VL or VH clones. A) 1 and 22- Ladder 1 kb; 2 to 11- VL Binders; 12 to 21- VL Internalized; 23- C^+; 24- C^-. B) 1 and 12- Ladder 1 kb; 2 to 11- VL Binders; 13 to 22- VL Internalized; 23- C^+; 24- C^-.

3.4.2. Expression and selection of specific antibodies

Once finalized the cloning of the VL and VH DNA (binders and internalized) into the pT7-PL vector, a high throughput screening was performed. A total of 528 clones, including VL and VH clones, were analyzed via ELISA assay to evaluate their expression and binding to the protein extract from Raji cells.

The figure below demonstrates the results for the VL clones (Fig. 18).

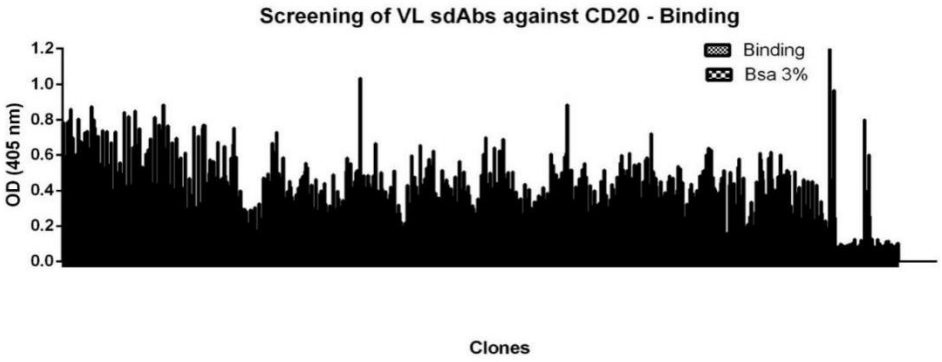


Figure 18 - **Relative binding activity of VL´s.** A total of 264 VL clones were randomly selected for ELISA to evaluate their binding activity against Raji protein extracts and BSA. HRP-conjugated anti-HA mAb was used as secondary antibody. Results were measured by optical density at 405 nm. A total of 34 clones (12.8%) with 2-3 times higher binding for Raji extracts than BSA background signal were isolated. Rabbit serum and anti-CD20 antibody were used as positive controls. BL21 extracts and SB medium were used as negative controls.

As shown in Figure 18, the clones that demonstrated having a stronger binding to the protein cell extract and expression were chosen, resulting in a binding percentage of 12.8 %.

Regarding the VH clones, 14 were selected, resulting in a binder percentage of 5.30% (Fig. 19).

For both selections, the expression levels were also considered and all selected clones demonstrated high expression levels.

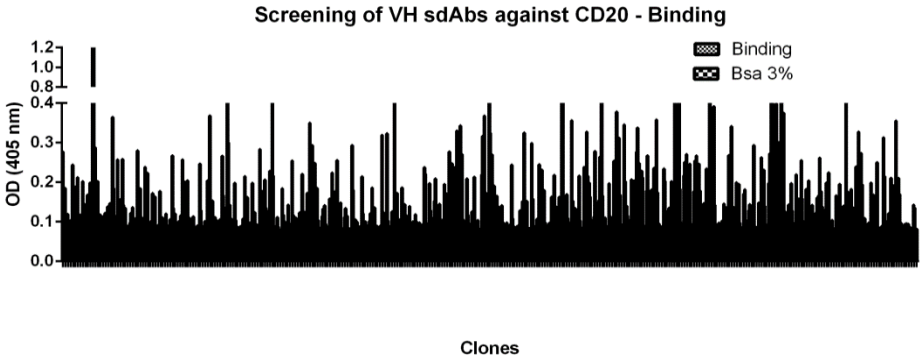


Figure 19 - **Relative binding activity of VH´s.** A total of 264 VH clones were randomly selected for ELISA to evaluate their binding activity against Raji protein extracts and BSA. HRP-conjugated anti-HA mAb was used as secondary antibody. Results were measured by optical density at 405 nm. A total of 14 clones (5.30 %) with 2-3 times higher binding than BSA background signal were isolated. BL21 extracts and SB medium were used as negative controls.

Each positive clone was tested again in ELISA assay with protein extract from Raji cells (data not shown). From this ELISA, 10 clones (9 VL binders and 1 VH binders) were selected and tested with Raji cells in order to evaluate the binding to the native CD20. The results of ELISA are presented in Figure 20.

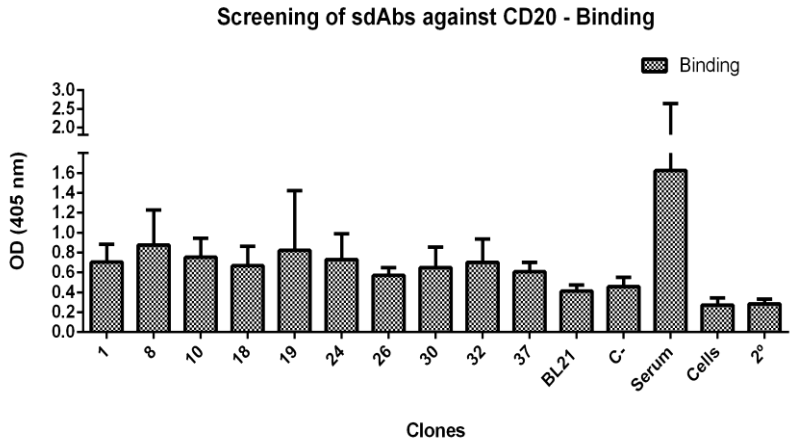


Figure 20 - **Relative binding activity of best VL and VH´s lead candidates identified.** A total of 10 clones (9 VL and 1 VH) were evaluated by cell ELISA for their binding activity against Raji cells. HRP-conjugated anti-HA mAb was used as secondary antibody. Rabbit serum was used as positive control and BL21 extracts and an irrelevant VL were used as negative control. Results were measured by optical density at 405 nm. Values represent mean±standard deviation of duplicates (n=2) for each condition, and similar results were obtained in two independent experiments.

The figure above indicates that the 10 clones are binding to the Raji cells. As a result of cells expressing the CD20 receptor in native conformation, we can conclude that the clones are recognizing it not only in linear conformation in the protein extract, but also in its native conformation present in lymphoma cells.

3.4.3. Sequencing of anti-CD20 clones

In order to evaluate the clone's profile, the 10 clones selected were sequenced and a homology analysis was performed by aligning the obtained amino acid sequences. The sequences analysis and respective homologies are present bellow.

As shown, the resulting sequences demonstrate that 9 clones belong to the same family of VLs, K2/j0. The remaining clone belongs to the VH3 family. Moreover, we can observe that the 9 VL clones are the same clone, since they have a homology of 100% (Fig. 21). Therefore, it seems that during the phage display there was a clone that was enrich during each panning.

Figure 21 - **Homology among amino acid sequences of selected clones.** Homology tree obtained from the homology analysis of selected the clones.

3.5.4. Confirmation of specificity for the CD20 receptor

By aligning the amino acid sequence obtained from sequencing, it was found that nine of the clones were equal, thus leaving only two clones (1 VL and 1 VH). The specificity of the two clones for the CD20 receptor was evaluated by Western Blot, one of which (VL antibody format) was shown to be specific to the CD20 receptor. The result for the VL clone is shown below (Fig. 22). No binding was observed for the VH clone and further studies are in progress to evaluate his binding.

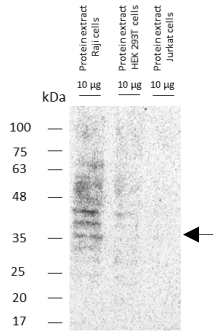


Figure 22 - **Western blot analysis of the total cell extract from the Raji, HEK 293T and Jurkat cells with the selected clone.**

The results of western blot show that the selected VL clone recognized proteins in the protein extract of the Raji cells, and did not recognize proteins in the cell extract of the HEK 293T or Jurkat cells. The recognized proteins' molecular weight is approximately 35 kDa, which means that the chosen clone seems to be highly specific to the CD20 receptor.

4. Conclusion and future perspectives

For decades, chemotherapy and radiotherapy had been the only treatments available for B cell lymphomas. However, with the discovery of Rituximab, new prospects in the field of immunotherapies became available. Throughout the years, the discovery of new antibodies for the CD20 receptor contributed to the development of new therapeutics and the quality of life of patients. Nevertheless, tumoral cells are always changing, making the search for better antibodies a never-ending process.

To fight the necessity of searching for better antibodies, the development of a new, single domain antibody, specific for the CD20 receptor, was the goal of this project. Our approach began with several immunizations of one rabbit, resulting in a strong immune response. Complex and highly diverse immune libraries were then constructed. Subtractive and positive selection, performed by phage display, resulted in a specific pool of sdAbs-phages. Binding activity studies and sequencing of the resulting clones of this pool allowed us to select for two clones (1 VL and 1 VH). These results demonstrate that phage display in cells, using positive and negative cells, is an efficient method when the goal is to select against only one target. Finally, we demonstrated that the VL clone was specific for the CD20 receptor and the most promising sdAb candidate to proceed for further studies. The approaches and methodologies implement in this project were important and determinant, as they lead us to the goal of the project.

The selected single domain antibody has many advantages compared to other antibodies because of its reduced size. Its reduced size allows the antibody to be used for several therapeutic applications related with B-cell malignancies. In the future, our single domain antibodies can be fused with an Fc region, leading to the recruitment of mechanisms in the immune system, such as apoptosis, complement activation, etc., to kill tumoral cells. The selected sdAbs could be also conjugated with a drug or toxin through a linker; and lastly, conjugated with a genetically modified T cell to form a chimeric antigen receptor (CAR-T cells). The applications of this sdAb could consist a future challenge. Furthermore, several in vitro studies (ex: epitope mapping, affinity measurements, cell activity, etc.) and *in vivo* studies related with the toxicity and efficiency of the sdAbs should be performed.

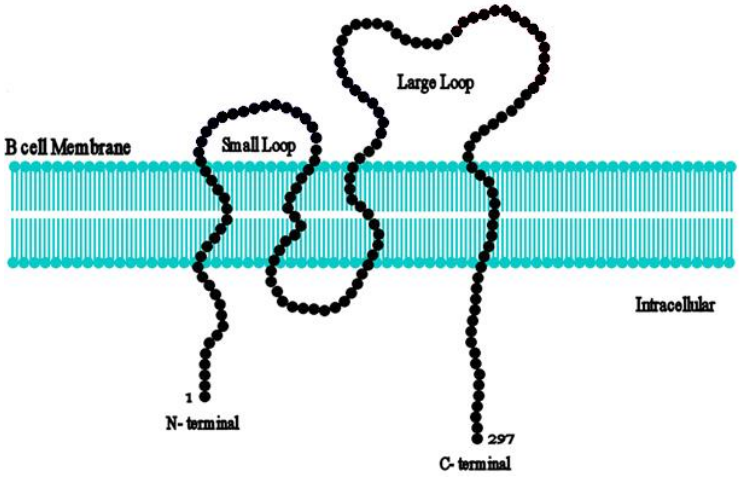
In conclusion, the results of this project are promising and can contribute to the development of several immunotherapies, namely for the B-cell malignancies, once the selected sdAb proves itself to be a favorable candidate to be used as a therapeutic agent. Furthermore, using these approaches, more candidates can be selected by opening doors to the development of new molecules for cancer immunotherapies and other emerging diseases.

5. References

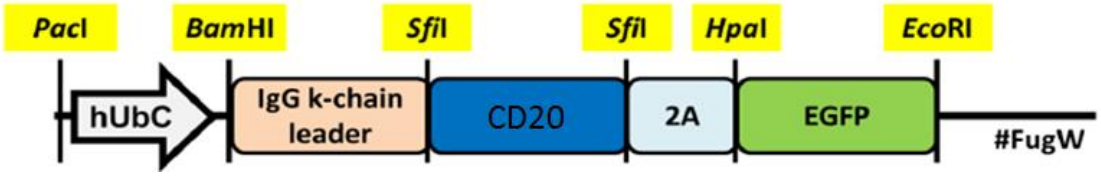
1. Siegel, R., Ma, J., Zou, Z. & Jemal, A. Cancer statistics, 2014. *CA. Cancer J. Clin.* **64,** 9–29 (2014).
2. American Cancer Society. Cancer facts & figures. *Am. Cancer Soc.* **8,** 1–34 (2013).
3. Shankland, K. R., Armitage, J. O. & Hancock, B. W. Non-Hodgkin lymphoma. *Lancet* **380,** 848–857 (2012).
4. Fisher, R. *et al.* Comparison of a standard regimen (CHOP) with three intensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. *N. Engl. J. Med.* **328,** 1002–1006 (1993).
5. Scott, A. M., Wolchok, J. D. & Old, L. J. Antibody therapy of cancer. *Nat. Rev. Cancer* **12,** 278–87 (2012).
6. Ecker, D. M., Jones, S. D. & Levine, H. L. The therapeutic monoclonal antibody market. *MAbs* **7,** 9–14 (2015).
7. Li, J. & Zhu, Z. Research and development of next generation of antibody-based therapeutics. *Acta Pharmacol. Sin.* **31,** 1198–1207 (2010).
8. Deans, J. P., Li, H. & Polyak, M. J. CD20-mediated apoptosis: signalling through lipid rafts. *Immunology* **107,** 176–182 (2002).
9. Lykken, J. M. & Tedder, T. F. The Tumor Microenvironment Regulates CD19 and CD20 Immunotherapy for Lymphoma. *Cancer J.* **21,** 351–356 (2015).
10. Press, O. W., Leonard, J. P., Coiffier, B., Levy, R. & Timmerman, J. Immunotherapy of non-Hodgkin ' s lymphomas. *Am. Soc. Hematol.* 221–240 (2001).
11. Pérez-Callejo, D., González-Rincón, J., Sánchez, A., Provencio, M. & Sánchez-Beato, M. Action and resistance of monoclonal CD20 antibodies therapy in B-cell Non-Hodgkin Lymphomas. *Cancer Treat. Rev.* **41,** 680–689 (2015).
12. Abbas, A., Listman, A. & Pillai, S. *Celular and molecular immunology*. (Elsevier, 2012).
13. Owen, J., Punt, J. & Stranford, J. *Kuby Immunology*. (W. H. Freeman and Company, 2013).
14. Male, D., Brostoff, J., Roth, D. & Roitt, I. *Immunology*. (Mosby Elsevier, 2007).
15. Klein, C. *et al.* Epitope interactions of monoclonal antibodies targeting CD20 and their relationship to functional properties. *MAbs* **5,** 22–33 (2013).
16. Storz, U. Rituximab: How approval history is reflected by a corresponding patent filing strategy. *MAbs* **6,** 820–837 (2014).
17. Baig, N., Taylor, R., Lindorfer, M., Church, A. & LaPlant, B. Induced resistance to Ofatumumab mediated cell clearance mechanisms, including complement dependent cytotoxicity, in Chronic Lymphocytic Leukemia. *J. Immunol.* **192,** 1620–1629 (2014).
18. Holliger, P. & Hudson, P. J. Engineered antibody fragments and the rise of single domains. *Nat. Biotechnol.* **23,** 1126–1136 (2005).
19. Hudson, P. J. Recombinant antibody constructs in cancer therapy. *Curr. Opin. Immunol.* **11,** 548–557 (1999).
20. Holt, L. J., Herring, C., Jespers, L. S., Woolven, B. P. & Tomlinson, I. M. Domain antibodies: Proteins for therapy. *Trends Biotechnol.* **21,** 484–490 (2003).
21. Dimitrov, D. & Marks, J. Therapeutic Antibodies: Current state and future trends - Is a paradigm change coming soon? *Methods Mol Biol.* **525,** (2009).
22. Kohler, G. & Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nat. Publ. Gr.* **253,** 495–497 (1975).
23. Tjandra, J. J., Ramadi, L. & McKenzie, I. F. Development of human anti-murine antibody (HAMA) response in patients. *Immunol. Cell Biol.* **68,** 367–376 (1990).
24. Almagro, J. C. & Fransson, J. Humanization of antibodies. *Front. Biosci.* 1619–1633 (2008). doi:10.1071/SHv11n1toc
25. Clackson, T., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. Making antibody fragments using phage display libraries. *Nature* **352,** 624–628 (1991).
26. Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science (80-.).* **228,** 1315–1317 (1985).
27. Barbas III, C. F., Burton, D., Scott, J. & Silverman, G. *Phage display: A laboratory manual*. (Cold Spring Harbor Laboratory Press, 2001).
28. Rader, C. & Barbas III, C. F. Phage display of combinatorial antibody libraries. *Curr. Opin. Biotechnol.* **8,** 503–508 (1997).
29. Vaughan, T. J. *et al.* Human antibodies with sub-nanomolar affinities isolated from a large non-immunized

phage display library. *Nat. Biotechnol.* **14,** 309–314 (1996).
30. Knappik, A. *et al.* Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J. Mol. Biol.* **296,** 57–86 (2000).
31. Aires, F. *et al.* Recombinat rabbit single-chain antibodies bind to the catalytic and C-terminal domains of HIV-1 integrase protein and strongly inhibit HIV-1 replication. *Biotechnol. Appl. Biochem.* **59,** 353–366 (2014).
32. Aires da Silva, F. & Gonçalves, J. Engineered rabbit antibody variable domains and uses thereof. (2008).
33. Cang, S., Mukhi, N., Wang, K. & Liu, D. Novel CD20 monoclonal antibodies for lymphoma therapy. *J. Hematol. Oncol.* **5,** 64 (2012).
34. Weiner, L. M., Surana, R. & Wang, S. Monoclonal antibodies: versatile platforms for cancer immunotherapy. *Nat. Rev. Immunol.* **10,** 317–27 (2010).
35. Aires da Silva, F., Corte-Real, S. & Goncalves, J. Recombinant Antibodies as Therapeutic Agents. *BioDrugs* **22,** 301–314 (2008).
36. Ruigrok, V. J. B., Levisson, M., Eppink, M. H. M., Smidt, H. & van der Oost, J. Alternative affinity tools: more attractive than antibodies? *Biochem. J.* **436,** 1–13 (2011).

Supplementary data



Supplementary figure 1- Molecular configuration of CD20 molecule. CD20 is a non-glycosylated phosphoprotein with four membranes spanning domains, the N - and C – terminal domains, which are intracytoplasmic and two additional extracellular loops. Adapted from Cang *et al*, 2012.[33]



Supplementary figure 2- **Construct used for transfection of HEK 293T cells.** This construct contains the sequence of the CD20 and sequence coding for GFP inserted into pFUGW vector.

VH 5' Sense Primers	
SDVH1-F	5' GGGCCCAGGCGGCCCAGTCGGTGGAGGAGTCCRGG 3'
SDVH2 -F	5' GGGCCCAGGCGGCCCAGTCGGTGAAGGAGTCCGAG 3'
SDVH3-F	5' GGGCCCAGGCGGCCCAGTCGYTGGAGGAGTCCGGG 3'
SDVH4-F	5' GGGCCCAGGCGGCCCAGSAGCAGCTGRTGGAGTCCGG 3'
VH 3' Antisense Primers	
SDG-R	5'CCTGGCCGGCCTGGCCACTAGTGACTGAYGGAGCCTTAGGTTGCCC 3'
VL 5' Sense Primers	
RSCVK1-F	5' GGGCCCAGGCGGCCGAGCTCGTGMTGACCCAGACTCCA 3'
RSCVK2-F	5'GGGCCCAGGCGGCCGAGCTCGATMTGACCCAGACTCCA 3'
RSCVK3-F	5' GGGCCCAGGCGGCCGAGCTCGTGATGACCCAGACTGAA 3'
RSCλ-F	5' GGGCCCAGGCGGCCGAGCTCGTGCTGACTCAGTCGCCCTC 3'
VL 3' Antisense Primer	
SDVLj10-B	5' CCTGGCCGGCCTGGCCTTTGATTTCCACATTGGTGCC 3'
SDVLj0-B	5' CCTGGCCGGCCTGGCCTAGGATCTCCAGCTCGGTCCC 3'
SDVL42j0-B	5'CCTGGCCGGCCTGGCCTTTGACSACCACCTCGGTCCC 3'
SDVLλ-B	5' CCTGGCCGGCCTGGCCGCCTGTGACGGTCAGCTGGGTCCC 3'

Supplementary figure 3- Sequence of oligonucleotides used for amplification of genes which codifying for rabbit VL and VH single domain antibodies.